Journal of Biometrics & Biostatistics
Yang et al., J Biom Biostat 2017, 8:3
DOI: 10.4172/2155-6180.1000356
Research Article
OMICS International

Nonparametric Estimation of Quantile and Quantile Density Function

Yang X[1]*, Hutson AD[1,2] and Wang D[3]

[1]*Department of Biostatistics, University at Buffalo, Buffalo, NY 14214, USA*
[2]*Roswell Park Cancer Institute, Buffalo, NY 14263, USA*
[3]*Department of Public Health and Preventive Medicine, SUNY Upstate Medical University, Syracuse, NY 13210, USA*

Abstract

In this article, we derive a new and unique method of estimating quantile and quantile density function, which is based on moments of fractional order statistics. A comparison of the proposed estimators is made with existing popular nonparametric quantile and quantile density estimators, in terms of mean squared error (MSE) for censored and uncensored data. Recommendations for the choice of quantile and/or quantile density estimators are given.

Keywords: Quantile function estimators; Quantile density function estimators; Order statistics; Kernel function

Introduction

The quantile function

$$Q(u) = \inf\{x : F(x) \ge u\}, \quad (1)$$

where F(.) is the cumulative distribution function (CDF) of a continuous random variable X and 0<u<1, is an alternative to the probability density function (PDF), the CDF and the characteristic function for describing a probability distribution. The estimation of Q(u) is of great interest, especially when one is unwilling to assume the distribution as parametric or when the underlying distribution is skewed.

The use of quantile function estimation has been around for decades in exploratory data analysis, statistical analysis, reliability and medical studies [1-11]. A more recent application can be found in Jeong and Fine [12] and Sankaran et al. [13] for competing risk models.

Many nonparametric estimators of the quantile function have been proposed and studied extensively. For uncensored data, the simplest method is the empirical quantile (EQ) estimator based on a single order statistic. It is a piecewise constant function that does not provide a useful quantile density function estimation. Details about advantages of smoothed quantile estimators can be found in Cheng and Parzen [7]. Numerous smoothed quantile function estimators have been introduced. Here, only the most representative ones are outlined. The most commonly used estimator is the linear interpolation of successive order statistics, which is employed in applications, for example, Q-Q plots and popular software packages such as SAS, BMDP, and MINITAB. Parzen [1] developed kernel smoothing of the EQ estimator, which is well known as the kernel quantile estimator. It has been extensively studied and analyzed [3-5,8]. More complete literature reviews on kernel-based quantile estimators can be found in Sheather and Marron [8] and Cheng and Parzen [7]. However, kernel quantile estimators in general are complicated and analytically intractable. Their performance in the sense of MSEs is very sensitive to the choice of bandwidth. In addition, the approximations to kernel estimators may violate the monotonicity requirement as described by Yang [5] and this issue is also shown in our simulation study and the real data application in Section 5 and Section 6. Generalized order statistics were considered as alternatives to sample quantiles by Harrell and Davis [14] and Kaigh and Lachenbruch [15]. Huang [16] proposed a modification of the Harrell-Davis (HD) estimator based on developing a weighting scheme through the use of the level crossing empirical distribution function.

In the presence of right-censored data, the product-limit quantile (PLQ) estimator proposed by Sander [17] and the general kernel smoothing version of PLQ estimator by Padgett [18], which share similar problems as their parallels for uncensored data, have gained the most popularity in the literature. More recently, Wang et al. [19] extended the HD quantile function estimator for censored data and proposed an exact bootstrap procedure for optimization in terms of MSE related criteria.

In the same way that the CDF can be differentiated to give the PDF, Parzen [1] and Jones [20] defined the derivative of Q(u) as the quantile density function. That is, q(u)=Q'(u). Common applications of q(u) include but are not limited to constructing the asymptotic confidence interval of sample quantiles, inference procedures based on linear rank statistics in Hettmansperger [21], and quantile density based approach in the location scale problem, see Eubank [22].

To estimate the quantile density function q(u), given either censored or uncensored data, two main approaches can be applied. One is the mathematical derivative of quantile estimators (if differentiable), the other is the reciprocal of density quantile function f(Q(u)) obtained by differentiating on both sides of the equation F(Q(u))=u. The former way is more advantageous over the latter in terms of efficiency; see Jones [20] for more information pertaining to the comparison of these two methods. In addition, kernel smoothing of the reciprocal of density quantile function has also been considered [23]. As mentioned before, PLQ and EQ estimators fail to provide useful quantile density function estimation. However, the linear interpolation of two successive order statistics is differentiable and the resulting quantile density function estimator is a histogram type estimator by Siddiqui [24]. The derivative of the kernel quantile estimator by Parzen [1] was introduced by Falk [25]. Xiang [26] proposed a natural derivative of the quantile estimator by Padgett [18]. More reviews on quantile density estimators can be found in Cheng and Parzen [7].

In this article we take a new and novel approach to quantile and quantile density function estimation based on estimating moments of

***Corresponding author:** Yang X, Department of Biostatistics, University at Buffalo, Buffalo, NY 14214-3000, USA, Tel: (716) 845-1300; E-mail: xyang37@buffalo.edu

Received May 26, 2017; **Accepted** June 16, 2017; **Published** June 20, 2017

Citation: Yang X, Hutson AD, Wang D (2017) Nonparametric Estimation of Quantile and Quantile Density Function. J Biom Biostat 8: 356. doi: 10.4172/2155-6180.1000356

Citation: Yang X, Hutson AD, Wang D (2017) Nonparametric Estimation of Quantile and Quantile Density Function. J Biom Biostat 8: 356. doi: 10.4172/2155-6180.1000356

fractional order statistics and solving a set of simultaneous equations pertaining to a series of moment expansions. We studied and compared the performance of our estimators with EQ estimator, PLQ estimator, the kernel smoothing of EQ and PLQ estimators, piecewise linear estimator and their corresponding quantile density estimators (if exist) for censored and uncensored data. The competing estimators were considered simply because they are commonly used for quantile and quantile density estimation. The advantages of our method are as follows: First, it does not require a selection of the optimal bandwidth and therefore can be more stable compared to the common kernel-based methods. Second, it at least outperforms the PL and the piecewise linear quantile estimators across all possible simulation parameters we considered in terms of MSEs and also appears to preserve the monotonicity of the quantile function curve. Third, the associated quantile density function estimator is shown to yield the smallest MSE among all quantile density estimators considered for both censored and uncensored data.

In Sections 2 and 3, we outline the existing methods of quantile and quantile density estimation considered in this investigation. In Section 4, our new quantile and quantile density estimators are introduced. The performance of our estimators is illustrated in terms of MSE by a Monte Carlo simulation study in Section 5. This is followed by an application of the switch life data reported by Nair [27] in Section 6. Recommendations of the choice of quantile, and/or quantile density estimators are summarized in Section 7.

Estimation of Q(u)

Let $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$ be the order statistics from a random sample with a continuous distribution $F(\cdot)$. And let $T_{(1)} \le T_{(2)} \le \cdots \le T_{(n)}$ be the order statistics corresponding to the i.i.d. sample of randomly right-censored times $T_1, T_2, \cdots, T_n$, and $\delta_{(1)} \le \delta_{(2)} \le \cdots \le {}_{(n)}$ are censoring indicators corresponding to the ordered $T_{(i)}$'s, respectively. A value of $\delta_{(i)}=1$.

indicates that $T_{(i)}$ is uncensored, while a value of $\delta_{(i)}=0$ indicates that $T_{(i)}$ is censored. The methods described below can be readily applied for uncensored data by setting all $\delta_{(i)}=1$.

Then the well-known PLQ estimator by Sander [17], is defined as

$$\hat{Q}_{PL}(u) = \inf\{x : 1 - \hat{S}_{PL}(x) >= u\}, \qquad (2)$$

where $\hat{S}_{PL}$ is the common PL estimator of the survival function by Kaplan and Meier [28]:

$$\hat{S}_{PL}(t) = \begin{cases} \prod_{T_{(j)} \le t} \left(\dfrac{n-j}{n-j+1} \right)^{\ddot{a}_{(j)}} \ t < T_{(n)} \\ 0,\ t \ge T_{(n)} \end{cases}. \qquad (3)$$

Defining $\hat{S}_{PL}(t) = 0$ for $t \ge T_{(n)}$ has been studied by Xiang [26] with respect to convergence properties for a class of kernel quantile function estimators, and was shown to provide a more technically suitable definition in term of the large sample theory as compared to setting $\hat{S}_{PL}(t)$ as undefined.

In the absence of censoring, $\hat{Q}_{PL}(u)$ reduces to the EQ estimator,

$$\hat{Q}_{EQ}(u) = \inf\{x : \hat{F}_E(x) >= u\}, \qquad (4)$$

where $\hat{F}_E(x)$ is the empirical distribution, $n^{-1}\sum_{i=1}^{n} I\left(X_{(i)} \le x\right)$.

The linear interpolation estimator of the quantile function given uncensored data is denoted as

$$\hat{Q}_L(u) = (1-\epsilon)X_{([n'u])} + \epsilon X_{([n'u]+1)}, \qquad (5)$$

where $\epsilon = n'u - [n'u]$ and $\epsilon = n'u - [n'u]$.

Padgett(1986) defined the kernel smoothing version of $\hat{Q}_{PL}(u)$ as

$$\hat{Q}_{KPL}(u) = h^{-1}\int_0^1 \hat{Q}_{PL}(u) K\left(\frac{t-u}{h}\right)dt = h^{-1}\sum_{i=1}^{n} T_{(i)} \int_{S_{i-1}}^{S_i} K\left(\frac{t-u}{h}\right)dt \qquad (6)$$

where S_i is the (3) at $T_{(i)}$ and K(.) is a symmetric kernel function. If no censoring, the estimator in eqn. (6) reduces to the general kernel smoothing of the EQ estimator by Parzen [1],

$$\hat{Q}_{KEQ}(u) = h^{-1}\sum_{i=1}^{n} X_{(i)} \int_{\frac{i-1}{n}}^{\frac{i}{n}} K\left(\frac{t-u}{h}\right)dt. \qquad (7)$$

Estimation of q(u)

Let, $\dfrac{i}{n'} \le u < \dfrac{i+1}{n'}, i = 1, 2, \ldots, n-1$ the first order derivative of $\hat{Q}_L(u)$ is:

$$\hat{q}_L(u) = n'\left(X_{(i+1)} - X_{(i)}\right) \qquad (8)$$

This is called the spacings of the sample, see Pyke [29,30], or a histogram type estimator by the fact that $\hat{q}_L(u) = 1/\hat{f}\left(\hat{Q}(u)\right)$, where $\hat{f}\left(\hat{Q}(u)\right)$ is the density quantile function estimator based on finite differences introduced by Siddiqui [24].

Again, the PLQ estimator and the EQ estimator do not have the corresponding quantile density estimators. The natural derivative of (6), $\hat{q}_{KPL}(u)$, was established by Xiang [26] as

$$\hat{q}_{KPL}(u) = -h^{-2}\sum_{i=1}^{n} T_{(i)} s_i K'\left(\frac{1-\hat{S}_{PL}\left(T_{(i)}\right) - t}{h}\right), \qquad (9)$$

where s_i denotes the jump of $\hat{S}_{PL}\left(T_{(i)}\right)$, that is,

$$s_i = \begin{cases} 1 - \hat{S}_{PL}\left(T_{(i)}\right), i = 1 \\ \hat{S}_{PL}\left(T_{(i-1)}\right) - \hat{S}_{PL}\left(T_{(i)}\right), i = 2, 3, \ldots, n \end{cases} \qquad (10)$$

When no censoring is present, eqn. (9) reduces to the kernel quantile density estimator $\hat{q}_{KEQ}(u)$, which is the complementary of Parzen [1] estimator in eqn. (7).

Fractional Order Statistic-Based Quantile and Quantile Density Estimator

In order to develop our new quantile and quantile density function estimators, we need to derive asymptotic expansions corresponding to the k^{th} non-central moment of the fractional order statistic, $E\left(X^k_{n'u:n}\right)$, where the fractional order statistic $X_{n'u:n}$ for an i.i.d. uniform sample jointly follow a particular Dirichlet process $U_{n'u:n}$; see Stigler [31]. Even though fractional order statistics do not exist in the empirical sense, their respective expectations may be calculated.

In deriving the expansion we assume that the first three derivatives of Q are bounded in a neighborhood of u and denote them as $Q' = Q'(u)$, $Q'' = Q''(u)$ and $Q''' = Q'''(u)$. We also assume, similar to the results pertaining to the i^{th} order statistics as in Section 3.1 and Section

3.2 of W. R. van Zwet (1964), that the expectation $E(X_{n'u:n})$ of the n' uth order statistic exists for some u and n. Then the asymptotic expansion is as the following:

Theorem: For large samples the first three non-central moments of $E\left(X^k_{n'u:n}\right) = \mu^k_{X_{n'u:n}}, k = 1, 2, 3$, are given by

$$\mu_{X_{n'u:n}} = Q(u) + \frac{u(1-u)Q''(u)}{2(n+2)} + O(n^{-2}), \quad (11)$$

$$\mu^2_{X_{n'u:n}} = Q(u)^2 + \frac{u(1-u)(2Q'(u)^2 + 2Q(u)Q''(u))}{2(n+2)} + O(n^{-2}), \quad (12)$$

$$\mu^3_{X_{n'u:n}} = Q(u)^3 + \frac{u(1-u)(6Q(u)Q'(u) + 3Q(u)^2Q''(u))}{2(n+2)} + O(n^{-2}), \quad (13)$$

Proof: First note that $X_{n'u:n} = Q(U_{n'u:n})$. Expanding $Q(U_{n'u:n})$ about $U_{n'u:n} = u$ and taking expectations of each term in the expansion based upon a Beta distribution $\beta(n'u, n'(1-u))$ proves the result. Note that the bounds in terms of Q''' is with respect to the remainder term of the expansion.

We may utilize equations 11-13 to find an approximation for Q(u) as a function o $E(X_{n'u:n})^k$, $k = 1,2,3$. Hence, we propose:

Definition: The moment quantile estimator, $\hat{Q}_M(u)$. is given by solving the set of simultaneous nonlinear equations

$$\hat{\mu}_{X_{n'u:n}} \approx Q(u) + \frac{u(1-u)Q''(u)}{2(n+2)}, \quad (14)$$

$$\hat{\mu}^2_{X_{n'u:n}} \approx Q(u)^2 + \frac{u(1-u)(2Q'(u)^2 + 2Q(u)Q''(u))}{2(n+2)}, \quad (15)$$

$$\hat{\mu}^3_{X_{n'u:n}} \approx Q(u)^3 + \frac{u(1-u)(6Q(u)Q'(u) + 3Q(u)^2Q''(u))}{2(n+2)}, \quad (16)$$

at a fixed u in terms of Q(u), Q'(u)and Q''(u)in order to estimate Q(u), where $\hat{\mu}^k_{X_{n'u:n}}$, $k = 1,2,3$, is given by the exact bootstrap kth moment estimator of the uth fraction order statistic by Hutson and Ernst [32],

$$\hat{\mu}^k_{X_{n'u:n}} = \sum_{i=1}^{n} X^k_{(i)} w_i. \quad (17)$$

The weight w_i is given as

$$w_i = B_{n'u,n'(1-u)}\left(\frac{i}{n}\right) - B_{n'u,n'(1-u)}\left(\frac{i-1}{n}\right) \quad (18)$$

for uncensored data and

$$w_i = B_{n'u,n'(1-u)}\left(1-\hat{S}_{PL}\left(T_{(i)}\right)\right) - B_{n'u,n'(1-u)}\left(1-\hat{S}_{PL}\left(T_{(i-1)}\right)\right) \quad (19)$$

for censored data in Wang et al. [19], where $B_{a,b}(x) = \int_0^x t^{a-1}(1-t)^{b-1}dt$ is the incomplete beta function. As an aside, estimates of Q'(u) and Q″(u) are also available as part of this process. In this investigation, we only interested in the performance of the first-order derivative of Q(u), i.e., the quantile density function q(u), and let's denote the quantile density estimator as $\hat{q}_M(u)$. In the later simulation, we will show that $\hat{q}_M(u)$ is at least more reliable than $\hat{q}_{KPL}(u)$, $\hat{q}_{KEQ}(u)$, and $\hat{q}_L(u)$ in terms of MSE.

If we are only interested in an estimate for Q(u) then the numerical solution with respect to Q(u) in terms of our system of equations 17-19 is relatively straightforward and reduces to solving the simple cubic equation

$$E^*(X_{n'u:n} - Q(u))^3 = 0, \quad (20)$$

with respect to Q(u), u fixed, where E^* denotes the exact bootstrap moment estimator of the quantity at (4.10). Alternatively, we can reformulate (20) and define it as follows:

Definition: The cubic quantile estimator, $\hat{Q}_C(u)$ is given by minimizing

$$\sum_{i=1}^{n} |X_{(i)} - Q(u)|^3 w_i, \quad (21)$$

with respect to Q(u), where the weights w_i are defined at (19) for censored data or at (18) for uncensored data. The associated quantile density function $\hat{q}_C(u)$ can still be estimated by taking the numerical derivative of an interpolated function from cubic quantile estimates at a proper number of u values. As motivated by this, an alternative way of estimating the moment quantile and quantile density functions can be interpolating a quantile function curve from the moment quantile estimates at a proper number of u values and taking the numerical first-order derivative of the interpolated function at interested u points, respectively. To distinguish with $\hat{Q}_M(u)$ and $\hat{q}_M(u)$, which are obtained by solving the equation system (14-16) simultaneously, we denote the alternatives as $\hat{Q}_{M.i}(u)$ and $\hat{q}_{M.i}(u)$. The simulation study in Section 5 shows that $\hat{Q}_M(u)$ and $\hat{Q}_{M.i}(u)$ yield basically the same quantile estimations. Even though $\hat{q}_M(u)$ and $\hat{q}_{M.i}(u)$ are different, they are both significantly better than $\hat{q}_{KPL}(u)$, $\hat{q}_{KEQ}(u)$, and $\hat{q}_L(u)$ in terms of MSE criteria.

Simulation Results

For the purpose of illustrating the behavior of our estimators, a straight forward simulation study was carried out for samples of size n=30, 50, 100 for Weibull distribution with the quantile function, $Q(u) = (-\log(1-u))^\theta$, θ=0.5,1,1.5 and across the standardized normal, exponential, extreme value, and logistic distributions. The censoring distribution was given as a uniform distribution uniform (0,T) with T=2,5 and uncensored case was also considered. We utilized fixed quantiles of u=0.25, 0.5 and 0.75. For each combination, 2,000 Monte Carlo simulations were utilized.

For censored data, a comparison of $\hat{Q}_{PL}, \hat{Q}_{KPL}, \hat{Q}_M, \hat{Q}_{M.i}$ and $\hat{Q}_C$ was made in terms of MSEs in Tables 1-4. For uncensored data, quantile estimators we considered are $\hat{Q}_{EQ}, \hat{Q}_L, \hat{Q}_{KEQ}, \hat{Q}_M, \hat{Q}_{M.i}$ and $\hat{Q}_C$, as shown in Tables 5 and 6. MSEs of the corresponding quantile density estimators, if exist, were also summarized in Tables 7-12. Epanechnikov kernel function, $0.75(1-u^2)I(|u| \le 1)$, was utilized here for the considered kernel estimators. Note that even though the triangular kernel is more commonly used for kernel quantile estimators in literature, see Padgett [18], Nair and Sankaran [33], and Soni [23], it fails to provide a useful derivative when calculating the $K^1(.)$ in the kernel quantile density estimators. The Epanechnikov kernel, which gives the optimal kernel, see Prakasa Rao [34], was studied by Soni [23] for comparing non-parametric quantile density estimators and our simulation showed that Q(u) estimation behaviors under Epanechnikov kernel and triangular kernel were quite close, which was also confirmed in the study of Soni [23]. For simplicity, only the result of the Epanechnikov kernel was presented here. Bandwidth

Citation: Yang X, Hutson AD, Wang D (2017) Nonparametric Estimation of Quantile and Quantile Density Function. J Biom Biostat 8: 356. doi: 10.4172/2155-6180.1000356

		***θ* = 0.5**			***θ* = 1**			***θ* = 1.5**		
n	**method**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**
30	$\hat{Q}_{PL}$	13.247	18.632	36.021	14.756	45.677	173.423	12.858	70.408	571.736
30	$\hat{Q}_{KPL}$	5.413	13.528	64.776	16.357	32.939	651.067	10.903	36.465	1200.693
30	$\hat{Q}_{M}$	11.117	16.527	72.228	14.665	44.066	527.885	14.576	66.225	1309.573
30	$\hat{Q}_{M.i}$	11.117	16.527	72.226	14.665	44.066	527.905	14.576	66.225	1309.574
30	$\hat{Q}_{C}$	9.492	14.319	22.232	13.03	33.949	107.402	13.362	47.344	389.307
50	$\hat{Q}_{PL}$	7.456	11.262	23.24	8.372	30.323	129.169	5.771	43.967	497.303
50	$\hat{Q}_{KPL}$	3.247	8.085	28.04	9.741	17.729	360.658	5.642	87.146	1218.629
50	$\hat{Q}_{M}$	6.433	10.034	24.182	8.021	28.461	318.258	6.907	45.937	1162.492
50	$\hat{Q}_{M.i}$	6.433	10.034	24.181	8.021	28.461	318.258	6.907	45.937	1162.487
50	$\hat{Q}_{C}$	5.697	8.746	16.202	7.177	24.194	54.206	6.263	36.878	240.124
100	$\hat{Q}_{PL}$	3.681	5.323	10.703	3.935	13.744	75.536	2.591	21.057	378.403
100	$\hat{Q}_{KPL}$	2.584	5.629	10.762	12.72	11.341	157.495	16.762	16.339	794.296
100	$\hat{Q}_{M}$	3.198	4.763	9.519	3.61	13.204	138.78	2.739	22.763	787.276
100	$\hat{Q}_{M.i}$	3.198	4.763	9.513	3.61	13.204	138.779	2.739	22.763	787.276
100	$\hat{Q}_{C}$	2.929	4.4	8.781	3.317	11.965	27.973	2.468	20.106	111.144

Table 1: Mean squared error of Q(×1000), T=2.

		Normal			**Exponential**			**EVD**			**Logistic**		
n	**method**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**
30	$\hat{Q}_{PL}$	61.4	56.205	80.77	14.972	52.288	164.457	52.086	90.331	246.574	183.639	151.429	336.384
30	$\hat{Q}_{KPL}$	38.166	20.288	58.405	12.275	36.824	525.884	39.379	42.463	511.026	168.324	97.933	302.179
30	$\hat{Q}_{M.i}$	59.2	65.269	69.11	14.62	46.719	514.84	49.776	75.557	414.817	203.765	175.495	343.174
30	$\hat{Q}_{M.i}$	59.2	65.269	69.107	14.62	46.719	514.832	49.776	75.557	414.815	203.765	175.495	343.174
30	$\hat{Q}_{C}$	47.021	41.931	51.586	12.758	35.976	102.749	40.936	57.167	129.728	154.008	106.64	165.655
50	$\hat{Q}_{PL}$	37.597	33.418	47.666	7.63	29.643	120.266	30.762	50.907	166.823	109.202	85.275	206.402
50	$\hat{Q}_{KPL}$	37.597	13.125	113.547	9.81	28.552	332.367	22.934	27.343	246.909	113.405	57.43	175.97
50	$\hat{Q}_{M}$	34.73	40.438	43.743	7.477	28.994	314.704	27.369	45.607	248.009	112.425	99.965	173.61
50	$\hat{Q}_{M.i}$	34.73	40.438	43.743	7.477	28.994	314.699	27.369	45.607	248.009	112.425	99.965	173.61
150	$\hat{Q}_{C}$	29.669	26.818	35.7	6.778	24.635	55.111	25.406	39.059	67.723	93.358	67.633	85.413
100	$\hat{Q}_{PL}$	18.866	14.112	23.3	3.874	12.969	75.43	15.696	23.431	95.473	53.101	36.898	105.636
100	$\hat{Q}_{KPL}$	15.029	7.782	37.868	14.438	11.071	297.886	11.924	56.015	121.979	63.618	40.671	89.408
100	$\hat{Q}_{M}$	17.434	19.795	21.456	3.554	12.626	140.229	13.324	21.08	96.522	50.908	48.445	71.062
100	$\hat{Q}_{M.i}$	17.434	19.795	21.455	3.554	12.626	140.229	13.324	21.08	96.522	50.908	48.445	71.062
1100	$\hat{Q}_{C}$	15.817	13.582	18.817	3.268	11.485	27.836	12.953	19.473	34.1	45.648	36.278	44.123

Table 2: Mean squared error of Q(×1000),Weibull,T = 5.

Citation: Yang X, Hutson AD, Wang D (2017) Nonparametric Estimation of Quantile and Quantile Density Function. J Biom Biostat 8: 356. doi: 10.4172/2155-6180.1000356

		θ = 0.5			**θ = 1**			**θ = 1.5**		
n	**method**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**
30	$\hat{Q}_{PL}$	10.859	13.27	22.34	14.146	37.672	142.768	11.943	64.733	457.705
30	$\hat{Q}_{KPL}$	3.597	8.835	17.584	13.142	23.019	120.039	15.396	105.723	324.724
30	$\hat{Q}_{M}$	9.309	12.372	20.406	13.291	42.257	164.929	13.767	105.495	414.619
30	$\hat{Q}_{M.i}$	9.309	12.372	20.408	13.291	42.257	164.929	13.767	105.495	414.619
30	$\hat{Q}_{C}$	8.127	10.686	17.804	11.777	35.497	124.617	12.512	79.644	278.786
50	$\hat{Q}_{PL}$	6.414	8	13.462	7.714	22.729	79.774	5.831	39.046	282.241
50	$\hat{Q}_{KPL}$	2.257	5.733	11.657	6.183	20.284	157.402	5.211	23.261	171.749
50	$\hat{Q}_{M}$	5.674	7.417	12.188	7.277	23.464	90.408	6.619	48.853	283.646
50	$\hat{Q}_{M.i}$	5.674	7.417	12.196	7.277	23.464	90.408	6.619	48.853	283.646
50	$\hat{Q}_{C}$	5.031	6.714	11.039	6.559	20.486	76.107	5.905	40.317	217.301
100	$\hat{Q}_{PL}$	3.245	4.175	6.825	3.701	11.242	38.589	2.492	17.671	133.052
100	$\hat{Q}_{KPL}$	1.329	2.931	5.81	3.637	12.15	35.236	2.394	14.675	254.84
100	$\hat{Q}_{M}$	2.837	3.85	6.242	3.393	10.937	38.892	2.675	19.171	154.735
100	$\hat{Q}_{M.i}$	2.837	3.85	6.239	3.393	10.937	38.892	2.675	19.171	154.735
100	$\hat{Q}_{C}$	2.62	3.617	5.779	3.127	10.109	34.754	2.385	17.257	128.109

Table 3: Mean squared error of Q(×1000), T = 5.

		Normal			**Exponential**			**EVD**			**Logistic**		
n	**method**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**
30	$\hat{Q}_{PL}$	61.4	53.732	66.443	14.609	41.123	141.977	51.487	78.876	190.112	183.639	142.001	223.828
30	$\hat{Q}_{KPL}$	37.694	16.412	42.86	10.783	21.891	113.932	94.098	31.522	137.074	193.813	57.821	201.164
30	$\hat{Q}_{M.i}$	59.191	63.566	66.781	13.215	45.305	156.931	49.248	78.867	199.736	203.257	181.754	241.536
30	$\hat{Q}_{M.i}$	59.191	63.566	66.781	13.215	45.305	156.931	49.248	78.867	199.736	203.257	181.754	241.536
30	$\hat{Q}_{C}$	46.989	40.512	52.121	11.473	36.475	121.961	40.547	60.993	148.648	153.82	109.909	173.994
50	$\hat{Q}_{PL}$	37.597	32.471	39.993	7.47	23.868	80.226	30.748	47.044	100.269	109.202	81.942	127.889
50	$\hat{Q}_{KPL}$	35.783	10.17	33.218	6.697	17.278	112.964	23.829	40.997	260.688	72.528	35.711	92.7
50	$\hat{Q}_{M}$	34.73	40.047	39.113	7.079	23.934	89.937	27.281	41.682	108.988	112.424	98.252	135.912
50	$\hat{Q}_{M.i}$	34.73	40.047	39.113	7.079	23.934	89.937	27.281	41.682	108.988	112.424	98.252	135.912
150	$\hat{Q}_{C}$	29.667	26.336	32.94	6.392	20.717	74.968	25.359	36.81	88.194	93.345	66.871	108.154
100	$\hat{Q}_{PL}$	18.866	14.212	20.272	3.654	11.216	37.26	15.703	21.681	47.944	53.101	36.621	62.186
100	$\hat{Q}_{KPL}$	16.904	4.886	25.055	3.358	10.189	54.867	10.007	9.765	38.076	65.78	18.422	57.201
100	$\hat{Q}_{M}$	17.434	19.655	18.3	3.376	10.709	36.847	13.393	19.592	47.157	50.909	47.86	60.004
100	$\hat{Q}_{M.i}$	17.434	19.655	18.3	3.376	10.709	36.848	13.393	19.592	47.156	50.909	47.86	60.004
1100	$\hat{Q}_{C}$	15.817	13.452	16.346	3.099	9.797	33.184	12.935	18.394	41.918	45.647	35.72	52.355

Table 4: Mean squared error of Q (×1000), Weibull, uncensored.

		θ = 0.5			θ = 1			θ = 1.5		
n	method	u = 0.25	u = 0.50	u = 0.75	u = 0.25	u = 0.50	u = 0.75	u = 0.25	u = 0.50	u = 0.75
30	$\hat{Q}_L$	9.62	11.277	18.28	11.899	33.147	109.119	9.352	57.895	380.881
30	$\hat{Q}_{EQ}$	10.7	12.471	20.634	14.213	36.217	128.025	11.423	61.913	437.653
30	$\hat{Q}_{KEQ}$	3.091	22.452	33.251	11.977	16.033	91.378	9.72	58.631	516.648
30	$\hat{Q}_M$	8.589	10.724	16.607	12.523	37.139	124.11	12.983	87.72	605.788
30	$\hat{Q}_{M.i}$	8.589	10.724	16.611	12.523	37.139	124.11	12.983	87.72	605.788
30	$\hat{Q}_C$	7.489	9.382	14.54	11.086	31.1	97.904	11.767	67.257	418.041
50	$\hat{Q}_L$	5.707	6.859	10.929	6.851	19.865	63.43	4.967	33.457	212.619
50	$\hat{Q}_{EQ}$	6.126	7.556	12.106	7.656	22.171	69.311	5.712	36.968	231.98
50	$\hat{Q}_{KEQ}$	1.947	4.382	8.281	6.367	12.328	49.203	5.31	40.178	231.919
50	$\hat{Q}_M$	5.169	6.47	9.996	6.963	20.852	66.823	6.426	42.851	280.134
50	$\hat{Q}_{M.i}$	5.169	6.47	9.993	6.963	20.852	66.823	6.426	42.851	280.134
50	$\hat{Q}_C$	4.633	5.869	9.074	6.245	18.353	56.985	5.653	35.693	217.612
100	$\hat{Q}_L$	2.955	3.553	5.385	3.464	9.998	30.3	2.363	16.108	97.287
100	$\hat{Q}_{EQ}$	3.092	3.874	5.97	3.617	10.569	33.835	2.414	17.005	111.129
100	$\hat{Q}_{KEQ}$	0.939	2.171	4.534	7.287	8.746	21.288	4.971	38.55	64.046
100	$\hat{Q}_M$	2.649	3.354	4.859	3.307	9.873	29.37	2.649	17.553	106.727
100	$\hat{Q}_{M.i}$	2.649	3.354	4.859	3.307	9.873	29.37	2.649	17.553	106.727
100	$\hat{Q}_C$	2.458	3.16	4.536	3.046	9.188	26.55	2.351	15.903	91.397

Table 5: Mean squared error of Q(×1000), uncensored.

		Normal			Exponential			EVD			Logistic		
n	method	u = 0.25	u = 0.50	u = 0.75	u = 0.25	u = 0.50	u = 0.75	u = 0.25	u = 0.50	u = 0.75	u = 0.25	u = 0.50	u = 0.75
30	$\hat{Q}_L$	60.84	48.796	59.372	12	34.299	107.76	49.799	66.839	146.884	184.702	126.487	186.103
30	$\hat{Q}_{EQ}$	61.4	52.766	63.226	14.046	37.151	121.397	51.428	74.166	158.894	183.639	138.679	196.871
30	$\hat{Q}_{KEQ}$	57.226	15.829	48.744	10.009	16.269	67.007	37.638	26.696	193.783	154.978	53.183	365.382
30	$\hat{Q}_M$	59.183	62.896	59.396	12.596	38.484	126.397	48.815	70.709	160.351	203.093	176.888	207.935
30	$\hat{Q}_{M.i}$	59.183	62.896	59.396	12.596	38.484	126.397	48.815	70.709	160.351	203.093	176.888	207.935
30	$\hat{Q}_C$	46.952	40.059	47.837	10.863	31.528	100.705	40.383	56.499	123.693	153.751	108.082	157.518
50	$\hat{Q}_L$	37.025	30.908	36.429	6.674	20.995	61.35	30.17	42.189	83.558	109.424	76.321	107.885
50	$\hat{Q}_{EQ}$	37.597	32.098	38.787	7.303	22.73	67.931	30.748	45.133	89.506	109.202	80.965	116.022
50	$\hat{Q}_{KEQ}$	28.952	9.827	31.515	7.042	26.832	81.404	21.237	16.118	112.673	71.916	31.587	82.484
50	$\hat{Q}_M$	34.731	39.671	34.94	6.875	21.648	64.674	27.272	40.574	85.399	112.424	97.44	109.596
50	$\hat{Q}_{M.i}$	34.731	39.671	34.94	6.875	21.648	64.674	27.272	40.574	85.399	112.424	97.44	109.596
50	$\hat{Q}_C$	29.667	26.029	30.048	6.172	18.916	55.55	25.325	36.138	71.398	93.345	66.339	91.895
100	$\hat{Q}_L$	18.621	15.299	18.813	3.403	10.172	30.263	15.356	20.735	41.36	53.474	39.596	55.149
100	$\hat{Q}_{EQ}$	18.866	14.595	19.944	3.53	10.749	33.872	15.704	20.715	44.078	53.101	36.926	58.093
100	$\hat{Q}_{KEQ}$	28.152	4.781	17.073	3.264	4.833	35.748	11.863	7.954	28.928	42.272	15.797	73.102
100	$\hat{Q}_M$	17.433	19.602	17.241	3.266	9.831	29.664	13.381	18.979	39.8	50.907	47.674	52.174
100	$\hat{Q}_{M.i}$	17.433	19.602	17.241	3.266	9.831	29.664	13.381	18.979	39.8	50.907	47.674	52.174
100	$\hat{Q}_C$	15.817	13.405	15.562	2.998	9.01	26.708	12.933	17.925	35.485	45.646	35.566	46.98

Table 6: Mean squared error of q(×1000), Weibull, T = 2.

		θ = 0.5			**θ = 1**			**θ = 1.5**		
n	**method**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**
30	$\hat{q}_{KPL}$	8149.28	21114.3	302430	5279.13	81367	10753.5	184243	30426.8	4376781
30	$\hat{q}_{M}$	190.06	138.632	1773.02	357.404	416.272	19065.2	554.12	1016.54	60623.1
30	$\hat{q}_{M.i}$	346.405	454.448	7033.58	375.337	1521.34	58892.9	472.144	2854.32	131391
30	$\hat{q}_{C}$	140.399	164.01	588.519	233.121	461.381	8392.33	334.335	808.027	35986.6
50	$\hat{q}_{KPL}$	5365.27	9166.75	79146.1	2799.42	31653.6	3.2E+08	8938081	6779.78	33467.5
50	$\hat{q}_{M}$	171.934	115.401	546.322	263.807	186.036	16876.7	334.348	589.967	71765.9
50	$\hat{q}_{M.i}$	218.037	234.301	3712.65	246.717	960.017	70666.7	258.313	1914.88	161146
50	$\hat{q}_{C}$	107.444	130.532	398.723	169.025	404.488	6448.35	220.841	596.617	32809.8
100	$\hat{q}_{KPL}$	3457.73	5540.73	39824.2	3866.26	76229.6	4700535	1754.15	96355.8	2.5E+07
100	$\hat{q}_{M}$	143.995	100.197	214.914	185.4	131.304	14898.8	180.215	272.313	78301
100	$\hat{q}_{M.i}$	138.296	141.534	546.366	149.814	426.399	68805.3	180.709	922.613	225776
100	$\hat{q}_{C}$	73.674	83.005	314.818	101.28	290.091	3769.66	110.307	600.577	27189.1

Table 7: Mean squared error of q(×1000), T = 2.

		Normal			**Exponential**			**EVD**			**Logistic**		
n	**method**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.75**	**u = 0.25**	**u = 0.25**	**u = 0.50**	**u = 0.75**
30	$\hat{q}_{KPL}$	32923.4	8200.43	389215	93756.5	141038	47665.2	136389	68124.8	37944.4	54747.8	278660	786858
30	$\hat{q}_{M}$	1253.95	711.916	1832.75	336.029	361.954	18553.9	770.943	820.779	20900.7	14222.9	1631.87	23260.3
30	$\hat{q}_{M.i}$	1504.34	2822.31	4729.17	357.54	1678.35	55976.9	1874.88	2758.74	57398.9	3792.06	7567.05	48062.3
30	$\hat{q}_{C}$	814.682	469.436	1176.31	234.564	433.048	8338.58	596.582	662.985	9350.68	3022.53	1230.23	10283.8
50	$\hat{q}_{KPL}$	1.8E+09	8264.46	21800.3	2806.18	9350.96	350462	111190	77323	2.8E+07	602767	177417	1255108
50	$\hat{q}_{M}$	1055.82	550.683	1276.25	273.726	176.808	16926.5	605.194	508.001	18993.3	14291.5	1229.24	18865.5
50	$\hat{q}_{M.i}$	960.359	1667.24	2538.14	249.059	742.928	69260	1087.41	1370.22	56219.3	2663.91	4035.12	43938.8
50	$\hat{q}_{C}$	594.841	343.731	904.337	168.349	416.507	6608.79	450.908	581.489	6998.28	1939.95	957.065	7112.76
100	$\hat{q}_{KPL}$	33449.9	8325.39	30866.9	5382.12	101981	51028.4	167707	10893.5	526053	434490	143301	618886
100	$\hat{q}_{M}$	917.332	365.142	977.522	187.784	125.629	14339	427.641	365.059	16055.1	14354	888.459	16108.1
100	$\hat{q}_{M.i}$	619.263	822.268	1107.62	149.805	392.538	64085.2	536.566	703.808	50992.1	1658.33	1804.27	33753.7
100	$\hat{q}_{C}$	375.799	215.821	659.105	100.431	278.377	3788.13	309.846	402.304	3641.83	1199.3	613.156	3578.73

Table 8: Mean squared error of q(×1000), Weibull, T = 5.

		θ = 0.5			**θ = 1**			**θ = 1.5**		
n	**method**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**
30	$\hat{q}_{KPL}$	3254.05	5038.94	6887.1	17748.6	7139.63	866352	20238.5	48852.4	2328631
30	$\hat{q}_{M}$	158.016	81.839	213.034	300.251	183.126	7409.52	482.898	513.333	33910.9
30	$\hat{q}_{M.i}$	281.619	245.544	463.94	292.757	782.814	5512.51	381.838	3165.88	24081.6
30	$\hat{q}_{C}$	110.409	107.756	341.551	194.985	516.728	2790.31	324.463	1675.97	9230.97
50	$\hat{q}_{KPL}$	3241.09	4910.1	69152.1	33010.5	168218	16064.3	169396	7966.03	431404
50	$\hat{q}_{M}$	144.37	83.057	204.728	226.47	145.642	7452.77	302.343	365.374	34197.4
50	$\hat{q}_{M.i}$	182.975	164.569	331.176	203.681	474.967	3657.56	221.098	1538.72	14659.9
50	$\hat{q}_{C}$	85.272	83.814	234.323	138.365	338.525	2387	195.897	976.238	6065.82
100	$\hat{q}_{KPL}$	3221.26	5363.8	3759.22	24398.1	73960.8	2697764	1631024	8019.2	58598
100	$\hat{q}_{M}$	122.228	73.692	174.536	168.461	111.735	7390.44	164.246	239.365	34434
100	$\hat{q}_{M.i}$	118.142	96.882	225.505	135.451	295.26	1856.85	166.299	714.2	9585.25
100	$\hat{q}_{C}$	61.637	56.189	151.94	89.949	197.94	1333.92	100.662	456.847	6206.19

Table 9: Mean squared error of q(×1000), T=5

Citation: Yang X, Hutson AD, Wang D (2017) Nonparametric Estimation of Quantile and Quantile Density Function. J Biom Biostat 8: 356. doi: 10.4172/2155-6180.1000356

		Normal			Exponential			EVD			Logistic		
n	**method**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**
30	$\hat{q}_{KPL}$	38545	8553.6	54397.2	54012.1	7139.28	596465	24249.9	11207.7	500538	57260.4	21867.5	2E+07
30	$\hat{q}_{M}$	1253.5	654.914	1272.82	280.413	173.871	7443.92	752.581	592.599	10547.3	14228.1	1636.36	14070.1
30	$\hat{q}_{M.i}$	1495.8	2643.93	1774.43	270.254	773.599	4735.6	1747.54	1974.06	5364.8	3781.01	8183.3	5797.39
30	$\hat{q}_{C}$	812.46	417.749	964.279	185.737	500.469	2634.03	576.176	743.603	3127.67	3018.33	1283.02	3744.31
50	$\hat{q}_{KPL}$	2.4E+07	8598.16	240537	392563	5429.96	17970.9	154250	2741767	31780.1	109331	22086.3	280650
50	$\hat{q}_{M}$	1055.96	513.579	1116.81	235.759	154.3	7444.73	599.108	453.36	10658.6	14292.3	1137.33	14236.6
50	$\hat{q}_{M.i}$	960.447	1641.38	1224.61	206.582	496.723	3758.39	1059.6	1060.58	4119.6	2650.51	3747.45	4777.37
50	$\hat{q}_{C}$	594.953	311.538	721.009	140.666	315.854	2450.21	442.924	478.437	2876.64	1939.01	877.011	3303.95
100	$\hat{q}_{KPL}$	131996	8640.7	29730	1009626	5460.92	20570.2	73217.3	11335.3	74369.3	666840	22209.8	690653
100	$\hat{q}_{M}$	917.377	350.549	941.464	164.87	109.83	7394.43	427.406	331.716	10603.5	14354	850.491	14350
100	$\hat{q}_{M.i}$	618.658	818.937	771.579	130.775	281.672	1825.94	528.506	577.48	2267.99	1658.73	1721.42	2664.37
100	$\hat{q}_{C}$	375.79	205.667	445.164	88.231	197.073	1307.94	309.365	324.425	1596.78	1199.19	551.284	1823.81

Table 10: Mean squared error of q(×1000), Weibull, uncensored.

		$\theta = 0.5$			$\theta = 1$			$\theta = 1.5$		
n	**method**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**
30	$\hat{q}_{L}$	1639.43	1395.49	3123.14	1921.12	4130.41	19918.9	1525.46	7543.69	76720.2
30	$\hat{q}_{KEQ}$	3272.54	2705.34	1062.83	359942	5781.29	19822.4	84394.5	283373	52233.3
30	$\hat{q}_{M}$	140.88	69.881	191.952	268.893	165.64	7502.06	437.996	474.492	34726.3
30	$\hat{q}_{M.i}$	247.23	214.761	314.455	253.467	574.395	3120.29	331.449	2256.15	22712.6
30	$\hat{q}_{C}$	96.881	89.108	243.226	177.204	378.479	2340.34	292.876	1232.07	15177.3
50	$\hat{q}_{L}$	1489.12	1381.23	3554.46	1729.66	4091.42	21371.5	1251.51	7266.19	75372.1
50	$\hat{q}_{KEQ}$	3248.72	5506.37	4378.57	24746.6	6820.24	59502.6	52402.6	169446	151363
50	$\hat{q}_{M}$	129.451	72.218	182.486	210.65	136.673	7440.84	277.427	341.458	34565.5
50	$\hat{q}_{M.i}$	165.322	148.291	250.924	186.508	384.238	2033.78	210.659	1188.87	11177.7
50	$\hat{q}_{C}$	74.817	69.331	173.157	125.635	272.133	1428.3	176.52	763.439	7662.2
100	$\hat{q}_{L}$	1649.7	1469.24	2995.56	1908.75	4229.72	17156.8	1311.57	7029.24	56393.5
100	$\hat{q}_{KEQ}$	3819.37	6093.39	3596.05	15641.8	7435.75	43767.3	21528.8	123924	82623.3
100	$\hat{q}_{M}$	110.794	61.735	163.644	158.097	104.196	7383.99	159.322	227.886	34432.2
100	$\hat{q}_{M.i}$	108.645	78.736	170.205	125.585	233.879	1168.04	164.729	571.317	5263.25
100	$\hat{q}_{C}$	54.822	46.171	113.761	84.617	159.967	823.231	97.201	379.635	3754.77

Table 11: Mean squared error of q(×1000),uncensored.

h for kernel estimators was chosen based on minimizing the bootstrap MSE, in which 300 bootstrap samples with replacement at each value of u for each sample size, distribution and censoring combination were used [35-41].

From Tables 1-6, we conclude for nonparametric quantile estimators that:

$\hat{Q}_{KPL}$ or $\hat{Q}_{KEQ}$ has the smallest MSE in the majority of cases except when u is large. But again, its performance is unstable or sensitive to the choice of h (e.g., A large MSE occurs in Table 1, n=100, θ=1, and u=0.25). In addition, the process of computing the optimal bandwidth based on minimizing MSE can be time-consuming.

When data are heavily censored and skewed, it seems that only $\hat{Q}_{C}$ and $\hat{Q}_{C}$ are good at tails. For example, in Tables 1 and 2, when u=0.75, MSEs of other estimators under Weibull, exponential, extreme value, and logistic distributions are at least twice as large as the MSE of $\hat{Q}_{C}$. A further discussion on this is mentioned in Section 6.

		Normal			Exponential			EVD			Logistic		
n	**method**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**	**u = 0.25**	**u = 0.50**	**u = 0.75**
30	$\hat{q}_L$	10326.1	6195.28	12227.3	1823.09	4070.72	19696.3	8110.1	7892.31	26530.6	31925.3	15847	35215.1
30	$\hat{q}_{KEQ}$	22989.5	8603.5	269793	43385.8	5784.64	28342.2	107459	11724.5	34224.9	242847	22600.6	60192.5
30	$\hat{q}_M$	1253.29	613.463	1212.83	261.043	162.594	7508.28	740.556	550.844	10626.8	14229.4	1528.65	14108.8
30	$\hat{q}_{M.i}$	1494.93	2603.39	1358.93	247.874	583.998	3299.42	1719.6	1652.31	3832	3801.55	6976.98	4136.69
30	$\hat{q}_C$	811.772	396.356	757.417	168.081	372.244	2302.28	560.056	593.754	2675.48	3014.17	1132.32	3021.49
50	$\hat{q}_L$	10002.9	6245.06	10720	1909.79	4085.35	17228.1	7830.5	8447.68	25472.6	35739	15406.1	29693.7
50	$\hat{q}_{KEQ}$	28853.1	8639.82	250605	21039.6	10344.1	287873	69511.4	11733.7	37261.5	96031.1	22701.2	92945.2
50	$\hat{q}_M$	1055.99	493.033	1065.56	213.744	139.433	7431.73	598.543	432.766	10609.9	14292.1	1101.03	14126
50	$\hat{q}_{M.i}$	960.538	1613.79	954.132	185.119	398.384	2051.36	1050.6	939.772	2359.12	2651.84	3552.06	2709.39
50	$\hat{q}_C$	595.079	299.404	555.958	129.118	259.409	1394.92	436.709	421.911	1727.29	1938.53	820.323	1977.23
100	$\hat{q}_L$	9909.31	6215.52	11217.7	1913.43	4137.78	16641.8	8307.05	8135.47	23534.7	29755.4	16260.1	31476.4
100	$\hat{q}_{KEQ}$	25708.3	8660.66	235815	912186	5738.09	21414.9	168318	11737.4	57342	173060	22781.7	76443
100	$\hat{q}_M$	917.376	346.513	917.305	153.383	102.544	7386.75	427.416	301.712	10600	14354	830.98	14304.1
100	$\hat{q}_{M.i}$	618.594	812.926	646.631	123.076	234.206	1212.79	534.512	512.189	1523.53	1659.02	1678.65	1810.09
100	$\hat{q}_C$	375.762	202.429	379.963	82.181	160.428	853.959	309.609	288.904	1027.73	1199.03	531.529	1203.32

Table 12: Mean squared error of q(1000), uncensored.

Among estimators without the bandwidth selection, behaves the best for both censored and uncensored cases. And $\hat{Q}_M$ and $\hat{Q}_{M.i}$ are almost the same given fixed u.

1. And from Tables 7-12, we conclude for nonparametric quantile density estimators that:

$\hat{q}_C$ Produces the smallest MSE in almost all cases.

$\hat{q}_{KPL}$ or $\hat{q}_{KEQ}$ yields the largest MSE and is substantially larger than other quantile density estimators

$\hat{q}_M$ is the second best quantile density estimator in terms of MSE, which also implies that it is better than $\hat{q}_{M.i}$.

Application

A real life test data set with n=40 mechanical switches by Nair [27] was used for the purpose of illustration. The observed data is: T=(1.151*, 1.667, 2.119*, 2.547,1.170*, 1.695, 2.135, 2.548, 1.248*, 1.710, 2.197, 2.738*, 1.331*, 1.955*, 2.199*, 2.794, 1.381*, 1.965, 2.227, 2.883*, 1.499 , 2.012*, 2.250*,2.883*, 1.508*, 2.051*, 2.254, 2.910, 1.543*, 2.076*, 2.261*, 3.015, 1.577*, 2.109, 2.349*, 3.017, 1.584*, 2.116*, 2.369, 3.793*), where ∗ denotes a censored.

Observation: There is 57.5% censoring or 24 censored observations in this dataset. The choice of bandwidth h was borrowed from Padgett (1986) based on minimizing the bootstrap MSE: h=0.28 if $0 < u \leq 0.25$, h=0.34 if $0.25 < u \leq 0.90$, and h=0.40 if $0.90 < u \leq 1$.

$\hat{Q}_{M.i}$ and $\hat{q}_{M.i}$ were not included here since they do not show any additional advantages compared with the moment quantile method and the cubic quantile method in Definition 4.1-2 as shown in the simulation study.

Figures 1 and 2 showed the estimates of quantile and quantile density for this example data. We see from Figure 1 that $\hat{Q}_{KPL}$, $\hat{Q}_M$,

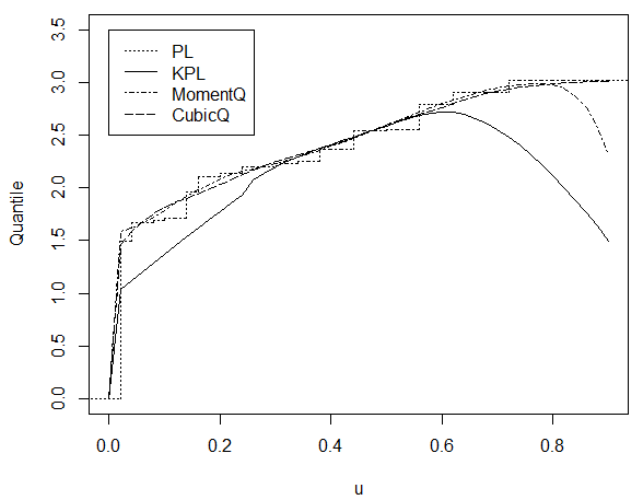

Figure 1: Quantile estimators for switch life data.

and $\hat{Q}_C$ do not differ much except at tails, and furthermore, only $\hat{Q}_{PL}$ and $\hat{Q}_C$ preserve monotonicity of quantile function curves for this data. These matches with our finding in simulation study that $\hat{Q}_{KPL}$ and $\hat{Q}_M$ can be away from the true quantile function for large values of u when data is heavily censored. Some techniques for correction at tails have already been explored and a brief review can be found in Soni et al. [23]. In addition, as we found in the simulation study, $\hat{q}_C$ performs the best among all quantile density estimators. The fluctuated curve of $\hat{q}_{KPL}$ in Figure 2 may also give a little hint about how bad the KPL method may perform in estimating quantile density functions.

Conclusion

In this article, we proposed three types of smooth quantile and

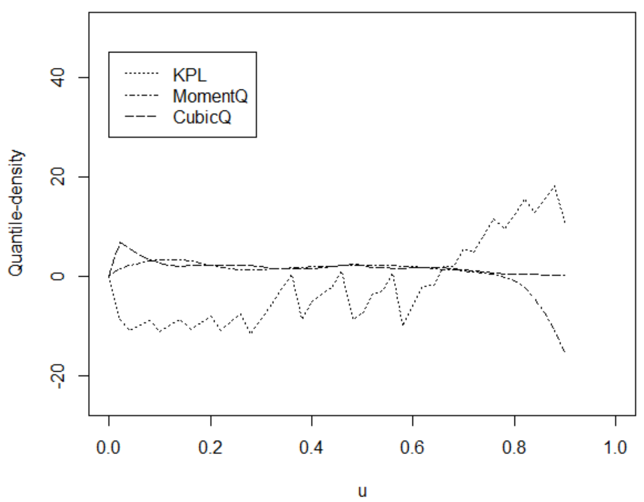

Figure 2: Quantile density estimators for switch life data.

quantile density function estimators. $\hat{Q}_M$ and $\hat{q}_M$ are better than $\hat{Q}_{M.i}$ and $\hat{q}_{M.i}$ in terms of MSEs and computational efficiency. But the cubic method, $\hat{Q}_C$ and $\hat{q}_C$, performs better than the two moment quantile methods mentioned above for both Q(u)and q(u) estimation, and also shows an obvious advantage in MSE to all the other nonparametric estimators, especially when it comes to quantile density function estimation.

In summary, if one is only interested in quantile estimation, $\hat{Q}_{KEQ}$ or $\hat{Q}_{KEQ}$, with modification at tails if in need, is a good choice. But if one prefers a more stable estimator, $\hat{Q}_L$ or $\hat{Q}_C$ may also be considered. For the estimation of quantile density function, $\hat{q}_C$ is clearly an optimal choice based on considerations of MSE, smoothness and simplicity. In addition, use $\hat{q}_L$ and $\hat{q}_{KPL}$ or $\hat{q}_{KEQ}$ with care for quantile density estimation since the bias can be extremely large compared to other alternatives.

Acknowledgement

This work was supported by Roswell Park Cancer Institute and National Cancer Institute (NCI) grant P30CA016056 and NRG Oncology Statistical and Data Management Center grant U10CA180822.

References

1. Parzen E (1979) Nonparametric statistical data modeling. Journal of the American Statistical Association 74: 105-121.
2. Parzen E (1993) Change PP plot and continuous sample quantile function. Communications in Statistics Theory and Methods 22: 3287-3304.
3. Reiss RD (1980) Estimation of quantiles in certain nonparametric models. Annals of Mathematical Statistics 8: 87-105.
4. Reiss RD (1989) Approximate Distributions of Order Statistics: With Applications to Nonparametric Statistics. Springer-Verlag, New York.
5. Yang S (1985) A smooth nonparametric estimator of a quantile function. Journal of the American Statistical Association 80: 1004-1011.
6. Cheng C (1995) The Bernstein Polynomial Estimator of a Smooth Quantile Function. Statistics & Probability Letters 24: 321-330.
7. Cheng C, Parzen E (1997) Unified Estimators of Quantile and Quantile Density Functions, Journal of Statistical Planning and Inference 59: 291-230.
8. Sheather SJ, Marron JS (1990) Kernel quantile estimators. Journal of the American Statistical Association 85: 410-416.
9. Hutson AD (1999) Calculating nonparametric confidence intervals for quantiles using fractional order statistics. Journal of Applied Statistics 26: 343-353.
10. Hutson AD (2001) Rational Spline Estimators of the Quantile Function. Communications and Statistics-Simulation and Computation 30: 377-390.
11. Hutson AD (2002) A Semiparametric Quantile Function Estimator for use in Bootstrap Estimation Procedures. Statistics and Computing 12: 331-338.
12. Jeong JH, Fine JP (2009) Parametric regression on cumulative incidence function. Biostatistics 8:184-196.
13. Sankaran PG, Nair NU, Sreedevi EP (2010) A quantile based test for comparing cumulative incidence functions of competing risks models. Statistics and Probability Letters 80: 886-891.
14. Harrell FE, Davis CE (1982) A new distribution-free quantile estimator. Biometrika 69: 635-640.
15. Kaigh WD, Lachenbruch PA (1982) A generalized quantile estimator. Comm. Statist A 11: 2217-2238.
16. Huang ML (2001) On a distribution-free quantile estimator, Computational Statistics Data Analysis 37: 477-486.
17. Sander J (1975) The Weak Convergence of Quantiles of the Product Limit Estimator, Technical, Report 5, Stanford University. Dept of Statistics.
18. Padgett (1986) A kernel-Type Estimator of a Quantile Function from right-censored data, Journal of the American Statistical Association 81: 215-222.
19. Wan D, Hutson AD, Gaile DP (2010) An exact bootstrap approach towards modification of the Harrell-Davis quantile function estimator for censored data. Journal of Nonparametric Statistics 22: 1039-1051.
20. Jones MC (1992) Estimating densities, quantiles, quantile densities and density quantiles. Annals of the Institute of Statistical Mathematics 44: 721-727.
21. Hettmansperger TP (1984) Statistical Inference Based on Ranks. Wiley, New York.
22. Eubank RL (1981) A density-quantile function approach to optimal spacings selection. Ann Statist 9: 494-500.
23. Soni, Dewanb I, Jain K (2012) Nonparametric estimation of quantile density function. Computational Statistics and Data Analysis.
24. Siddiqui MM (1960) Distribution of quantiles in samples from a bivariate population. J Res Nat Bureau Standards Section B 64: 145-150.
25. Falk M (1986) On the estimation of the quantile density function. Statist Probab Lett 4: 69973.
26. Xiang X (1994) A Law of the Logarithm for Kernel Quantile Density Estimators. The Annals of Probability 22: 1078-1091.
27. Nair VN (1984) Confidence Bands for Survival Functions With Censored Data: A Comparative Study. Technometrics 26: 265-275.
28. Kaplan EL, Meier P (1958) Nonparametric estimation from incomplete observations. Journal of the American Statistical Association 52: 457-481.
29. Pyke R (1965) Spacings. Journal of the Royal Statistical Society Ser B 27: 395-449.
30. Pyke R (1972).Spacings Revisited. Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, I, 417-427.
31. Stigler SM (1977) Fractional Order Statistics with Applications. American Statistical Society 72: 544-550.
32. Hutson AD, Ernst MD (2000) The Exact Bootstrap Mean and Variance of an Lestimator, Journal of the Royal Statistical Society–Series B 62: 89-94.
33. Nair NU, Sankaran PG (2009) Quantile based reliability analysis. Communications in Statistics Theory and Methods 38: 222-232.
34. Prakasa Rao BLS (1983) Nonparametric Functional Estimation. Academic Press, NewYork.
35. Cheng (2002) Almost-sure uniform error bounds of general smooth estimator's of quantile density functions, Statistics and Probability Letters 59: 183-194.
36. Hutson AD (1999) Calculating nonparametric confidence intervals for quantiles using fractional order statistics, Journal of Applied Statistics 26: 343-353.
37. Hutson AD (2001) Rational Spline Estimators of the Quantile Function. Communications and Statistics-Simulation and Computation 30: 377-390.

38. Hutson AD (2002) A Semiparametric Quantile Function Estimator for use in Bootstrap Estimation Procedures. Statistics and Computing 12: 331-338.

39. Srivastava DK, Mudholkar GS, Mudholkar A (1992) Assessing the significance of difference between two quick estimates of location. Journal of Applied Statistics 19: 405-416.

40. Vanzwet WR (1964) Convex Transformation of Random Variables. Mathematical Centre Tracts, Amsterdam.

41. Wang D, Hutson AD (2011) A fractional order statistic towards defining a smooth quantile function for discrete data. J Statist Plann Inference 141: 3142-3150.

Citation: Yang X, Hutson AD, Wang D (2017) Nonparametric Estimation of Quantile and Quantile Density Function. J Biom Biostat 8: 356. doi: 10.4172/2155-6180.1000356